

**IMAGE RETRIEVAL METHOD
BASED ON COMBINATION OF COLOR AND IMAGE FEATURES**

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to an image retrieval method, and more particularly, to an image retrieval method based on color and texture 5 describers that are significant in human visual perception. The present application is based on Korean Application No. 2001-1130, filed on January 9, 2001, which is incorporated herein by reference.

2. Description of the Related Art

In recent years, a sudden increase in multimedia data has facilitated 10 research in the fields of data retrieval and databases for efficient multimedia data management. In particular, research in the field of image retrieval has focused on an alternative to existing text-based image retrieval methods or further advanced image retrieval methods. For text-based data retrieval, a great effort needs to be made to retrieve the desired data from a large-scale 15 database, and it is unsuitable for partial-feature-based retrieval or quantitative data retrieval. To address these drawbacks, approaches have been suggested

for a retrieval method using numeric feature descriptors which well represent the features of the multimedia data.

- Color and texture features, which are dominant characteristics in human visual perception, can be expressed by feature descriptors. According
5 to a conventional image retrieval method, an image database in which data images are indexed by a feature descriptor is built up, and desired images are retrieved from the image database based on the feature descriptor. An example of a conventional color feature based image retrieval method includes extracting color vectors, quantizing the color vectors, getting a representative
10 color vector from the quantized color vectors, and retrieving the image data using the representative color vector. A conventional texture feature based image retrieval method uses the co-occurrence matrix, the Gaussian Markov Random Field (GMRF) models, etc. A recent method of texture feature extraction uses a Gabor filter that enables a multi-channel approach suitable
15 for the human perception mechanism of visual images.

- In the conventional image retrieval methods, color and texture based data retrievals are separately carried out, and color and texture features are extracted from the entire image. For this reason, the conventional methods cannot be applied to an image having a plurality of objects therein or
20 expressed in multiple colors and textures. Either the color or the texture features are available for retrieving particular image data. Thus, there is a problem of incorrect image retrieval results.

To solve these problems, another conventional image retrieval method retrieves a target image in the vector space formed for both color and texture vector components. Here, the vector space is formed through simple summations of the color and texture vector components, thereby increasing 5 dimensionality of the vector space. Thus, there is a need to redefine a distance function to the vector space with the increased dimensionality and to calculate a weighting factor for each component, which is relatively burdensome. The computations are highly complex and the results may be incorrect.

SUMMARY OF THE INVENTION

10 To solve the above-described problems, it is an object of the present invention to provide an image retrieval method based on the combination of color and texture features, and with improved performance.

15 To achieve the object of the present invention, there is provided a method for retrieving a data image similar to a query image, using a feature distance calculated by combining color and texture distances by considering human visual perception attributes.

In one embodiment, the present invention provides a method for retrieving a data image similar to a query image, in an image database containing a plurality of data images, the method comprising: (a) calculating 20 color and texture distances between a query image and each data image in the image database; (b) weighting the calculated color and texture distances with predetermined weighting factors; (c) calculating a feature distance between the

query image and each data image by combining the weighted color and texture distances by considering human visual perception attributes; and (d) determining the data image similar to the query image using the feature distance. Preferably, before step (a), the image retrieval method further

- 5 comprises: (pa-1) segmenting the query image and each data image into a plurality of regions using color features; and (pa-2) determining sample regions in the query image and each data image for extraction of texture features.

Preferably, step (a) comprises: (a1) generating color vectors of the
10 plurality of regions using the color features and calculating color distances; and (a2) generating texture vectors of the plurality of regions using the texture features and calculating texture distances.

Preferably, step (pa-1) comprises quantizing color vectors of the query image and each data image. This step of quantizing the color vectors may
15 comprise: (pa-1-1) performing peer group filtering on the query image and each data image for noise removal and smoothing effects; and (pa-1-2) clustering filtered pixel values of the query image and each data image using the generalized Lloyd algorithm.

Preferably, the image retrieval method further comprises: defining a J-value indicating color uniformity in each pixel of the query image and each
20 data image which have undergone quantization; storing the J-value in each pixel of the query image and each data image to obtain J-images; segmenting

each J-image into a plurality of regions by a predetermined segmentation method; repeating the segmentation of each J-image to obtain a map of over-segmented regions for each J-image; and obtaining a final map for each J-image by merging the over-segmented regions based on correlation of color;

- 5 Preferably, step (pa-1-2) comprises applying a predetermined algorithm to increase the number of resulting clusters or to merge the resulting clusters.

In the image retrieval method, the color features are expressed by a color feature descriptor $f_c(I_k)$ with a representative color vector and the percentage of the representative color vector for each region. The color feature descriptor $f_c(I_k)$ may be expressed by:

$$f_c(I_k) = \{(\bar{c}_{k1}, p_{k1}), (\bar{c}_{k2}, p_{k2}), \dots, (\bar{c}_{kN_k}, p_{kN_k})\}$$

where k is a positive integer indicating the serial number of each region, \bar{c}_k is the i-th representative color vector of the k-th region ($i=1,2,\dots,N$), p_{ki} is the percentage of the i-th representative color vector in the k-th region, and N_k is the number of the representative color vectors in the k-th region.

Preferably, the image retrieval method further comprises indexing a feature vector space by a representative color and its percentage in each segmented region. This step of indexing the feature vector space may 20 comprise: assigning representative colors to grid points in a color space having a grid structure; and storing the result of the assignment as a table in a

database.

In the image retrieval method, the color distance between each color vector of each region of the query image and each color vector of each region of each data image may be calculated using:

$$d_c(I_1, I_2) = \sum_{i=1}^{N_1} P_{1i}^2 + \sum_{i=1}^{N_2} P_{2i}^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2\alpha_{1i,2j} P_{1i}P_{2j}$$

where, when T_d is the maximum distance by which similarity of two colors is determined, α is a predetermined scaling coefficient, $d_{max} = \alpha T_d$, and d_y is the Euclidean distance $\|c_i - c_j\|$ between two color vectors c_i and c_j ,

$$\alpha_y = 1 - \frac{d_y}{d_{max}}$$
 if $d_y \leq T_d$, and $\alpha_y = 0$ if $d_y > T_d$.

In the image retrieval method, preferably, step (a2) uses the Gabor function. In step (a2), the texture vectors of the plurality of regions may be generated using the Gabor function having N frequency channels and M orientation channels, where N and M are predetermined positive integers.

In the image retrieval method, the texture features are expressed by a texture feature descriptor $f_t(I_k)$:

$$f_t(I_k) = \{(m_{k11}, \sigma_{k11}), (m_{k12}, \sigma_{k12}), \dots, (m_{k1M}, \sigma_{k1M}), (m_{k21}, \sigma_{k21}), \dots, (m_{ky}, \sigma_{ky}), \dots, (m_{kNM}, \sigma_{kNM})\}$$

where m_{kj} is the mean value of the pixel values of the i-th frequency channel and j-th orientation channel for region I_k , and σ_{kj} is the deviation of the pixel values of the i-th frequency channel and the j-th orientation channel for region I_k .

- 5 In step (a2), the texture distance between each color vector of each region of the query image and each color vector of each region of each data image may be calculated using:

$$d_t(I_1, I_2) = \sum_{i=1}^N \sum_{j=1}^M \left| \frac{m_{ij} - m_{2j}}{\sigma(m_j)} \right| + \sum_{i=1}^N \sum_{j=1}^M \left| \frac{\sigma_{1j} - \sigma_{2j}}{\sigma(\sigma_j)} \right|.$$

- Preferably, step (pa-2) comprises: (pb-1) obtaining a grid map of the query image and each data image; and (pb-2) obtaining a texture sample of a 10 desired size for each region based on the grid map. In this case, step (pb-1) may comprise: (pb-1-1) taking a rectangle $M(i, j)$ having largest side lengths for a sample region, where $0 \leq i \leq n$ and $0 \leq j \leq m$; (pb-1-2) dividing the rectangle $M(i, j)$ into sub-rectangles each having a $l \times l$ size; and (pb-1-3) obtaining a grid map for the sample region using a predetermined function

$$M_{\text{grid}}(i, j) = \begin{cases} 1 & \text{if } p_g(x, y) \subset I_k \text{ for } \forall (x, y) \\ 0 & \text{otherwise} \end{cases}$$

- 15 which outputs 1 if all the divided grids belong to the sample region, and otherwise, outputs 0, the predetermined function expressed by:

where $0 \leq i \leq \left\lceil \frac{n}{l} \right\rceil - 1$, $0 \leq j \leq \left\lceil \frac{m}{l} \right\rceil - 1$, $0 \leq x \leq l - 1$, $0 \leq y \leq l - 1$, and $p_y(x, y)$

is a point that belongs to rectangle $M(il + x, jl + y)$. Preferably, after step (pb-1-3), the image retrieval method further comprises: (pb-1-4) transforming the grid map obtained in step (pb-1-3) to a grid distance map by repeating

5 computations with:

$$\{M_{\text{grid}}^d(i, j)\}_0 = M_{\text{grid}}(i, j), \text{ and}$$

$$\{M_{\text{grid}}^d(i, j)\}_n = \min(\{M_{\text{grid}}^d(i, j - 1)\}_{n-1}, \{M_{\text{grid}}^d(i - 1, j)\}_{n-1}, \{M_{\text{grid}}^d(i, j + 1)\}_{n-1}, \{M_{\text{grid}}^d(i + 1, j)\}_{n-1}) + 1,$$

$$\text{if } \{M_{\text{grid}}^d(i, j)\}_{n-1} = \{M_{\text{grid}}^d(i, j)\}_n, M_{\text{grid}}^d(i, j) = \{M_{\text{grid}}^d(i, j)\}_n, \text{ and}$$

(pb-1-5) performing region growing in every direction from a seed of the grid

10 map, which has the largest distance in the grid and is expressed by $(a, b) = \arg \max_{(i, j)} \{M_{\text{grid}}^d(i, j)\}$, to satisfy the relation $M_{\text{grid}}^d(i, j) > 0$ until the maximum area rectangle is extracted.

Preferably, the image retrieval method further comprises fitting the maximum area rectangle extracted in step (pb-1-5) to the desired size of the texture sample by wrapping. Alternatively, the image retrieval method may further comprise fitting the maximum area rectangle extracted in step (pb-1-5) to the desired size of the texture sample by mirroring.

In the image retrieval method, step (b) may comprise: (b-1) placing the color and texture distances in 2-dimensional vector spaces defined by the

respective color and texture distances and their predetermined weighting factors; and (b-2) projecting the result of the projection onto the 2-dimensional vector spaces onto a 1-dimensional distance space based on the human visual perception mechanism.

- 5 In the image retrieval method, preferably, the predetermined weighting factor to the color distance is determined based on the distribution of representative colors. More preferably, the predetermined weighting factor to the color distance is determined by:

$$\omega_c = 1 - \sum_{i=1}^N P_{qi} \log_{10}\left(\frac{1}{P_{qi}}\right)$$

- where P_{qi} is the percentage of the i -th representative color of the query image
10 region.

Preferably, the texture distance is defined by:

$$d_t(I_q, I_1) = \alpha^{\frac{A(s)}{A(I_1)} + (count(I_1)-1)} d_t(I_q, I_1)$$

- where I_q denotes the query image or each region of the query image, s denotes the sample region, $A(\cdot)$ denotes the area of the sample region of a
15 desired size, $count(\cdot)$ is the number of wrappings done to obtain the desired size of sample region, and α is a constant.

Preferably, the predetermined weighting factor to the texture distance is determined based on the area of an initial sample region extracted from the

query image and the area of a sample region extracted from each data image.

More preferably, the predetermined weighting factor to the texture distance is determined by:

$$\omega_i = \frac{1}{\alpha^{\left(\frac{A(s)}{A(I_q)} + (count(I_q) - 1)\right)}}$$

- 5 where I_q denotes the query image or each region of the query image, s denotes the sample region, $A(\cdot)$ denotes the area of the sample region of a desired size, $count(\cdot)$ is the number of wrappings done to obtain the desired size of sample region, and α is a constant.

- Preferably, the image retrieval method further comprises, before step
 10 (b-1), normalizing each of the color and texture distances. In this case, normalizing each of the color and texture distances may be performed using Gaussian normalization. Preferably, normalizing each of the color and texture distances comprises: under the assumption that sequence v_k is a Gaussian sequence, performing normalization within the range of [-1, 1] using

15 $v_{m,k} = \frac{v_{m,k} - \mu_k}{3\sigma_k}$ based on the mean value μ_k and deviation σ_k of sequence

v_k ; and mapping the result of the normalization into the range of [0, 1] using

$$v'_{m,k} = \frac{v_{m,k} + 1}{2}.$$

Preferably, normalizing each of the color and texture distances comprises normalizing the texture distance by updating the mean value μ_t and

deviation σ_t of the texture distances by excluding a largest texture distance

until the condition of $k \times \frac{\mu_c}{\sigma_c} \leq \frac{\mu_t}{\sigma_t}$, where k is a constant, is satisfied.

In the image retrieval method, step (b-2) may comprise projecting the color and texture distances onto a 1-dimensional distance space using:

5 $d(I_q, I_1) = W_c d_c(I_q, I_1) \left(1 + \frac{2}{\pi} \tan^{-1} \frac{d'_t(I_q, I_1)}{d_c(I_q, I_1)} \right) + W_t d'_t(I_q, I_1) \left(1 + \frac{2}{\pi} \tan^{-1} \frac{d_c(I_q, I_1)}{d'_t(I_q, I_1)} \right)$
where $W_c = \frac{\sigma_c}{\sigma_c + \sigma_t}$, and $W_t = \frac{\sigma_t}{\sigma_c + \sigma_t}$.

BRIEF DESCRIPTION OF THE DRAWINGS

The above object and advantages of the present invention will become more apparent by describing in detail preferred embodiments thereof with

10 reference to the attached drawings in which:

FIG. 1 is a flowchart illustrating the major steps of a preferred embodiment of an image retrieval method according to the present invention;

FIG. 2 is a flowchart illustrating an example of indexing data images of an image database by color feature;

15 FIG. 3A shows an example of an image segmented according to its color features;

FIG. 3B shows an image expressed with quantized color vectors for the image of FIG. 3A;

FIG. 3C shows a J-image;

FIG. 3D shows an image after segmentation is completed;

FIG. 4 is a flowchart illustrating indexing data images of an image database by texture feature;

5 FIG. 5 shows an example of the spatial frequency response equal to half of the 2-D Gabor function;

FIG. 6 is a flowchart of an example of taking a texture sample from an image region;

FIG. 7 illustrates a feature of the present invention;

10 FIG. 8A shows examples of data image regions for a given query image region, which are retrieved by combining color and texture features; and

FIG. 8B shows examples of data image regions for a given query image region, which are retrieved based on only color feature; and

15 FIG. 8C shows examples of data image regions for a given query image region, which are retrieved by considering texture feature on the candidate data image regions having significant color distances.

DETAILED DESCRIPTION OF THE INVENTION

In an image retrieval method according to the present invention, a data
20 image having similar features as those of a query image is extracted from an image database in which data images are indexed by feature descriptors. FIG.

1 is a flowchart illustrating the major steps of a preferred embodiment of the
image retrieval method according to the present invention. Referring to FIG.
1, the image retrieval method according to the present invention involves
calculating color and texture distances between a query image and each of the
5 data images stored in the image database (Step 102).

In the present embodiment, it is assumed that each of the data images
stored in the image database is segmented into a plurality of regions according
to color and/or texture features, and an image region that a user wants to
search for in the image database exists in one data image segmented into a
10 plurality of image regions. That image region which is searched for is referred
to as a query image region. In the image retrieval method according to the
preferred embodiment of the present invention, a data image similar to the
query image region is retrieved.

First, a step of extracting color (feature) vectors for the query image
15 and data images and calculating color distances will be described. FIG. 2
shows an example of indexing the image database by color feature. Referring
to FIG. 2, color vectors for the query image and data images are quantized
(Step 202). FIG. 3A shows an example of an image segmented according to
its color features, and FIG. 3B shows an image expressed with the quantized
20 color vectors for the image of FIG. 3A. Next, peer group filtering is
performed on the images with the quantized color vectors for noise removal
and smoothing effects (Step 204). Peer group filtering is described in a Patent

Application PCT/KR00/00090 filed on February 3, 2000 by the present applicant and incorporated by reference herein. Next, filtered pixel values are clustered using the generalized Lloyd algorithm (GLA) (Step 206). Although the number of clusters is fixed for the GLA, a predetermined algorithm can be

- 5 applied to increase the number of clusters or merge the clusters. Next, a J-value as an index of color uniformity in each pixel after the quantization is defined (Step 208). For the images expressed by the quantized color values, the J-value indicating pixel color uniformity is stored in each of the pixels to obtain a J-image (Step 210). FIG. 3C shows the J-image obtained in Step 210.
- 10 The J-image is segmented into a plurality of regions by a predetermined method (Step 212). The J-value based image segmentation is disclosed in a Patent Application PCT/KR00/00248 filed on March 22, 2000 by the present applicant and incorporated by reference herein. A map of over-segmented regions is obtained by iterating segmentation (Step 214). Next, a final map is
- 15 obtained by merging segmented regions based on correlation of color (Step 216), thereby completing the image segmentation. FIG. 3D shows an image after the segmentation is completed.

Next, a feature vector space is indexed by the representative color and its percentage in each segmented region (Step 218). For example, a color feature describer $f_c(I_k)$ of each segmented region is expressed using a representative color vector and the percentage of the representative color vector:

$$f_c(I_k) = \{\bar{c}_{k1}, p_{k1}\}, (\bar{c}_{k2}, p_{k2}), \dots, (\bar{c}_{kN_k}, p_{kN_k})\} \quad \dots(1)$$

where k is a positive integer indicating the serial number of each region, \bar{c}_k is the i -th representative color vector ($i = 1, 2, \dots, N$) of the k -th region, p_k is the percentage of the i -th representative color vector, and N_k is the number of
5 the representative color vectors in the k -th region. Then, the representative colors are assigned to grid points in a color space having a grid structure. Next, the indexed results are stored as a table in a database, thereby completing indexing of the feature vector space by the color vectors.

The distance between color vectors of a query image region and a data
10 image region is expressed by:

$$d_c(I_1, I_2) = \sum_{i=1}^{N_1} p_{1i}^2 + \sum_{i=1}^{N_2} p_{2i}^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2\alpha_{ij} p_{1i} p_{2j} \quad \dots(2)$$

In formula (2), when T_d is the maximum distance by which similarity of two colors is determined, α is a predetermined scaling coefficient, $d_{max} = \alpha T_d$, and d_y is the Euclidean distance $\|c_i - c_j\|$ between two color vectors c_i ,
15 and c_j , $\alpha_y = 1 - \frac{d_y}{d_{max}}$ if $d_y \leq T_d$, and $\alpha_y = 0$ if $d_y > T_d$.

Next, a step of extracting texture (feature) vectors and calculating texture distances will be described. FIG. 4 is a flowchart illustrating indexing an image database by texture feature.

The reason for the wide use of multi-channel filtering in recent years as a texture analysis approach is that the similarity of this multi-channel filtering to human visual perception is evidenced by psychophysical and psychophysiological experiments. The mammalian visual tissue has a

- 5 perception mechanism for tuning visual signals into narrow band frequencies and orientations. In other words, the visual perception mechanism of humans can be explained with a plurality of channels each having a bandpass filter with particular frequency and orientation angle domains. The important factor in multi-channel filtering texture analysis is to define the functional
10 characteristics of the channels and the number of channels. In addition, the extraction of correct texture features from the image passed through the filters, the relation between each channel, and the appropriate combination of texture features from different channels are also important considerations. The most effective function for modelling texture features is the Gabor function.
15 Referring to FIG. 4, images are filtered by Gabor filtering (Step 402). A 2-dimensional (2-D) Gabor function used for the Gabor filtering of images can be defined as a complex sinusoidal function, which is a modulation of a 2-D Gaussian function the spectrum of which is shifted in the frequency domain.

This function is expressed by:

$$h(x, y) = g'(x', y') \exp[2\pi f(U_x + V_y)] \quad \dots(3)$$

20

$$(x', y') = (x \cos \phi + y \sin \phi - x \sin \phi + y \cos \phi) \quad \dots(4)$$

2010-10-26 10:42:02

$$g(x,y) = \frac{1}{2\pi\lambda\delta^2} \exp[-(x^2/\lambda^2 + y^2)/2\delta^2] \quad \dots(5)$$

$$H(\mu, \nu) = \exp\{-2\pi^2\delta^2[(\mu' - U')^2 + (\nu' - V')^2]\} \quad \dots(6)$$

where λ is the aspect ratio, and δ^2 is variance.

Supposing that the radial bandwidth is B and the orientation bandwidth

- 5 is Ω , the radial bandwidth B and the orientation bandwidth Ω of the Gabor filter are determined such that about half of the spatial-frequency response for each channel filter is the same to reduce redundancy in the Gabor filter. FIG. 5 shows an example of the spatial frequency response equal to half of the 2-D Gabor function. Referring to FIG. 5, a first channel 502 has the same scale as
10 a second channel 504, but different orientation. The first channel 502 has a different scale from a third channel 506, but the same orientation. Gabor filtering for an image is implemented by convolving the image with the Gabor function, thereby outputting different images through filters having different orientations and scales.

- 15 The mean value and variance of the pixel values of the filtered image are calculated (Step 404). The mean value and variance of the pixel values obtained by convolution of the image with the Gabor filter represent texture features of the image. Next, the feature vector space is indexed by the mean value and variance of the pixel values (Step 406). The image segmented into a

plurality of regions is expressed using the Gabor filter having N frequency channels and M orientation channels:

$$f_i(I_k) = \{(m_{k11}, \sigma_{k11}), (m_{k12}, \sigma_{k12}), \dots, (m_{k1M}, \sigma_{k1M}), (m_{k21}, \sigma_{k21}), \dots, (m_{kj}, \sigma_{kj}), \dots, (m_{kNM}, \sigma_{kNM})\} \quad \dots(7)$$

where m_{kj} is the mean value of the pixel values of the i-th frequency channel

and j-th orientation channel for image region I_k , and σ_{kj} is the deviation of

- 5 the pixel values of the i-th frequency channel and the j-th orientation channel for image region I_k .

A distance function for two image regions based on their texture vectors is expressed as:

$$d_t(I_1, I_2) = \sum_{i=1}^N \sum_{j=1}^M \left| \frac{m_{1j} - m_{2j}}{\sigma(m_j)} \right| + \sum_{i=1}^N \sum_{j=1}^M \left| \frac{\sigma_{1j} - \sigma_{2j}}{\sigma(\sigma_j)} \right|. \quad \dots(8)$$

- According to the present invention, texture features are extracted from
- 10 an image region having an arbitrary shape. An image region selected by a user has an irregular shape. Color information is irrelevant to the shape of the image region because it is attainable from the number and distribution of color pixels in the image region. Supposing that texture is defined as an irradiance pattern that can be discriminated by its proportion in the spatial frequency
 - 15 domain, the texture feature is affected by pixel patterns and the shape of image regions. It is not easy to apply the Gabor filter that outputs the convolution of an image with the Gabor function to an image having an arbitrary shape.

When an image is retrieved through the Gabor filter using texture samples having different sizes, matching efficiency is poor compared with an image retrieval based on the same-sized samples. For this reason, in the present embodiment, there is a need to take rectangular samples from image regions

- 5 having an arbitrary shape, which are the same in size and represent the texture feature of corresponding regions. In this sampling process, the rectangular samples should be large enough to include the irradiance pattern of the pixels, such that accurate texture information is extracted from the relevant image regions.

- 10 FIG. 6 shows an example of taking a texture sample from an image region. In this step, the texture sample is obtained using the gray image and a map of a sample region. Referring to FIG. 6, a grid map for the sample region is obtained to take a texture sample (Step 602). In particular, a rectangle $M(i, j)$ (where, $0 \leq i \leq n$ and $0 \leq j \leq m$), which has largest side lengths, is taken from the sample region. Next, the rectangle $M(i, j)$ is divided into sub-rectangles each having a $l \times l$ size. Then, a grid map for the sample region is made with the following predetermined function, which outputs 1 if all the divided grids belong to the sample region, and otherwise outputs 0:
- 15

$$M_{\text{grd}}(i, j) = \begin{cases} 1 & \text{if } p_y(x, y) \subset I_k \text{ for } \forall (x, y) \\ 0 & \text{otherwise} \end{cases} \quad \dots(9)$$

where $0 \leq i \leq \left[\frac{n}{l} \right] - 1$, $0 \leq j \leq \left[\frac{m}{l} \right] - 1$, $0 \leq x \leq l - 1$, $0 \leq y \leq l - 1$, and $p_y(x, y)$

is a point that belongs to rectangle $M(il + x, jl + y)$.

Next, a maximum area rectangle is extracted (Step 604). The grid map obtained in Step 602 is transformed into a grid distance map by repeated

5 computations with:

$$\{M_{\text{grid}}^d(i, j)\}_0 = M_{\text{grid}}(i, j), \text{ and} \quad \dots(10)$$

$$\{M_{\text{grid}}^d(i, j)\}_n = \min(\{M_{\text{grid}}^d(i, j - 1)\}_{n-1}, \{M_{\text{grid}}^d(i - 1, j)\}_{n-1}, \{M_{\text{grid}}^d(i, j + 1)\}_{n-1}, \{M_{\text{grid}}^d(i + 1, j)\}_{n-1}) + 1,$$

$$\text{if } \{M_{\text{grid}}^d(i, j)\}_{n-1} = \{M_{\text{grid}}^d(i, j)\}_n, M_{\text{grid}}^d(i, j) = \{M_{\text{grid}}^d(i, j)\}_n. \quad \dots(11)$$

10 From a seed of the grid distance map, $(a, b) = \arg \max_{(i, j)} \{M_{\text{grid}}^d(i, j)\}$, which has the largest distance in the map, region growing is carried out in every direction to satisfy the relation $M_{\text{grid}}^d(i, j) > 0$ until the maximum area rectangle is extracted.

Next, a texture sample having a desired size is obtained (Step 606).

15 The maximum area rectangle obtained in Step 604 is fit to the desired size of the texture sample by wrapping. Alternatively, the maximum area rectangle may be fit to the desired size of the texture sample by mirroring. In this case, orientation of the texture should be maintained constant after the mirroring.

- Referring to FIG. 1, after Step 102, color and texture distances are weighted with predetermined weighting factors (Step 104). In the present embodiment of the image retrieval method according to the present invention, color and texture vectors of the data images are compared with those of a
- 5 query image in the respective vector spaces to separately calculate the color and texture distances. Referring to FIG. 7, the calculated color and texture distances are projected onto 2-D vector spaces with weighting factors. The feature values on the 2-D vector spaces are projected onto a 1-D distance space by considering human perceptual attributes.
- 10 In particular, the distance between the color vector of a query image region and the color vector of a data image is defined by the distance function of formula (2) above. A weighting factor applied in this step is determined according to the color distribution. For example, if a dominant color is present, color reliability is great. On the contrary, if a plurality of
- 15 representative colors are present in similar proportion without a dominant color, color reliability is low. The color distribution is expressed by the entropy function. When the color distribution is uniform over an image region, the entropy function has a larger value. When the color distribution is concentrated on a small portion of the image region, the entropy function has a
- 20 smaller value. Based on this, a weighting factor to the color feature is defined as:

$$\omega_c = 1 - \sum_{i=1}^N P_{qi} \log_{10}\left(\frac{1}{P_{qi}}\right) \quad \dots(12)$$

where p_{q_i} is the percentage of the i-th representative color in the query image region. The color-weighting factor ω_c is set to 0.1 if $\omega_c < 0.1$.

For the texture feature vectors of the query image region and the data image, texture distance is determined using the function having formula (8), in the same manner as for the color distance. When the method described with reference to FIG. 6 is applied to take a texture sample, texture reliability may be low if the texture sample is small. Although the size of the extracted texture sample can be fit to a desired size by wrapping, unsatisfactory texture features may result with increased number of wrappings, due to discontinuity at the wrapping boundaries. Accordingly, in the present embodiment, each initial texture distance $d_t(I_q, I_1)$ is multiplied by a weighting factor.

$$d'_t(I_q, I_1) = \alpha^{\frac{A(s)}{A(I_1)} + (count(I_1)-1)} d_t(I_q, I_1) \quad \dots(13)$$

where I_q denotes the query image or query image region, s denotes the texture sample of a desired size, $A(\cdot)$ denotes the area of the texture sample of a desired size, $count(\cdot)$ is the number of wrappings done to obtain the desired size of texture sample, and α is a constant.

In a similar way, a weighting factor to the texture feature is defined as:

$$\omega_t = \frac{1}{\alpha^{\frac{A(s)}{A(I_q)} + (count(I_q)-1)}} \quad \dots(14)$$

where $A(\cdot)$ denotes the area of the texture sample of a desired size.

- Next, the distance between the query image region and data image is calculated by the combination of the color and texture distances weighted by considering human perception attributes (Step 106). As shown in FIG. 7, the color and texture distances weighted with the respective weighting factors in
5 Step 104 are projected onto a 1-D distance space through a projector that is based on the human visual perception attributes. To combine the two discrete features, each weighted color and texture distance needs to normalize into similar scales. The normalization is performed using Gaussian normalization in the present embodiment. Under the assumption that sequence v_k is a
10 Gaussian sequence, normalization is performed within the range of [-1, 1] by formula (15) using the mean value μ_k and deviation σ_k of sequence v_k with a 99% confidence factor. Next, the result is normalized within the range [0,1] by formula (16).

$$v_{m,k} = \frac{v_{m,k} - \mu_k}{3\sigma_k} \quad \dots(15)$$

$$15 \quad v'_{m,k} = \frac{v_{m,k} + 1}{2} \quad \dots(16)$$

According to this normalization method using formula (16), values less than 0 or greater than 1 are set to 0 or 1, respectively. This method is effective to prevent the range of the values to be normalized from being narrow due to the presence of extremely large or small values.

For the color distance function, the distance between color vectors is meaningless if a data image has no similar color component to the query image, and a deviation in color distance is small between candidate data images. On the contrary, the distance between texture vectors has a great

- 5 deviation between candidate data images, and thus the mean value μ_t and deviation σ_t of texture distances is updated by excluding the largest texture distance until the following condition (17) is satisfied, and then the texture distance is normalized using the updated mean value μ_t and deviation σ_t :

$$k \times \frac{\mu_c}{\sigma_c} \leq \frac{\mu_t}{\sigma_t} \quad \dots(17)$$

- 10 where k is a constant.

The normalized color and texture vector distances are combined by the following function (18) by weighting the color and texture vector distances with weighting factors:

$$d(I_q, I_1) = W_c d_c(I_q, I_1) \left(1 + \frac{2}{\pi} \tan^{-1} \frac{d'_t(I_q, I_1)}{d_c(I_q, I_1)} \right) + W_t d'_t(I_q, I_1) \left(1 + \frac{2}{\pi} \tan^{-1} \frac{d_c(I_q, I_1)}{d'_t(I_q, I_1)} \right) \dots(18)$$

15

$$\text{where } W_c = \frac{\sigma_c}{\sigma_c + \sigma_t}, \text{ and } W_t = \frac{\sigma_t}{\sigma_c + \sigma_t}.$$

Next, a data image similar to the query image is determined using the combined color and texture distances (Step 108). As can be inferred from the function (18), when one of the texture and color distances is much smaller

than the other, each of the texture and color distances is weighted with different weighting factors, which reflects well the human visual perception mechanism in which similarity is perceived by one influential texture or color feature. In the image retrieval method according to the present invention, each
5 multidimensional feature vector is projected onto respective 2-D vector spaces each having distance and weighting factor as elements, and then onto a 1-D distance space that defines a combined distance of discrete feature vectors. Accordingly, overall computations are relatively simple. In addition, the projector used in the image retrieval method according to the present
10 invention, which uses a small number of weighting factors which adaptively vary depending on, for example, the distribution of feature vectors, enables color and texture vector based image retrieval according to human visual perception.

As described above, according to the present invention, color and
15 texture features extracted from each data image region segmented by color are separately projected as feature vectors onto respective 2-D vector spaces each having elements of weighting factor and corresponding feature vector distance from a query image, and in turn projected onto a 1-D distance space based on human visual perception attributes. Thus, the image retrieval method
20 according to the present invention closely matches the human visual perception mechanism, and overall computations are relatively simple.

To verify performance of the image retrieval method according to the present invention, simulation tests were carried out using 100 image samples (384x256 or 256x384) designed by Corel Corporation. After segmenting the image samples into 2456 image regions in total using a color based 5 segmentation algorithm, color and texture vectors were extracted from each segmented region using color vector and texture vector extraction algorithms. Here, color feature vectors were extracted in the LUV color space, and texture feature vectors were extracted by applying a Gabor filter having 4 frequency channels and 6 orientation channels to texture samples of a 64x64 size. As a 10 result, forty-eight texture vector components were extracted from each region.

FIG. 8A shows examples of data image regions for a given query image region, which are retrieved by combining color and texture features. FIG. 8B shows examples of data image regions for a given query image region, which are retrieved based on only the color feature. FIG. 8C shows 15 examples of data image regions for a given query image region, which are retrieved by considering the texture feature on the candidate data image regions having significant color distances. Referring to FIGS. 8A through 8C, data images similar to the respective query images (here, fish's bodies) are retrieved based on data image regions, i.e., fish's body regions, in which both 20 the color and the texture of the scales are distinct. The resultant data images are satisfactory in view of human perception attributes. Thus, according to the

present invention, high-level features of an image can be effectively expressed by the appropriate combination of low-level features, color and texture.

- In the image retrieval method according to the present invention described above, data images similar to a query image can be retrieved based
- 5 on the human visual perception mechanism by combining both color and texture features extracted from image regions. In particular, this image region based retrieval enables accurate retrieval of many objects and many kinds of information from a single image. The distribution of lower-level features, color and texture, in the 2-D vector spaces is projected onto a 1-D distance
- 10 space, and thus higher-level features can be effectively expressed with a small number of weighting factors and reduced vector computations.

- The image retrieval method according to the present invention may be embodied as a program stored on a computer usable medium, including but not limited to storage media such as magnetic storage media, optically readable media, and carrier waves. A functional program, code and code segments, used to implement the present invention can be derived by a skilled computer programmer from the description of the invention contained herein.